OPTIMIZING THE THERMAL TRANSPORT PROPERTIES OF SINGLE LAYER (2D) TRANSITION METAL DICHALCOGENIDES (TMD)

Yenal KARAASLAN [1], Haluk YAPICIOĞLU [2,*], Cem SEVİK [1]

[1] Department of Mechanical Engineering, Faculty of Engineering, Eskişehir Technical University, Eskişehir, Turkey
[2] Department of Industrial Engineering, Faculty of Engineering, Eskişehir Technical University, Eskişehir, Turkey

ABSTRACT

In order to characterize thermal dependent physical properties of materials, potentially to be used in technological applications, an accurate interatomic-potential parameter set is a must. In general, conjugate-gradient methods and more recently, metaheuristics such as genetic algorithms are employed in determining these interatomic potentials, however, especially the use of metaheuristics specifically designed for optimization of real valued problems such as particle swarm and evaluation strategies are limited in the mentioned problem. In addition, some of these parameters are conflicting in nature, for which multi objective optimization procedures have a great potential for better understanding of these conflicts. In this respect, we aim to present a widely used interatomic potential parameter set, the Stillinger–Weber potential, obtained through three different optimization methods (particle swarm optimization, PSO, covariance matrix adaptation evolution strategies, CMA-ES, and non-dominated sorting genetic algorithm, NSGA-III) for two-dimensional materials MoS_2, WS_2, WSe_2, and $MoSe_2$. These two-dimensional transition metal dichalcogenides are considered as a case mainly due to their potential in a variety of promising technologies for next generation flexible and low-power nanoelectronics, (such as photonics, valleytronics, sensing, energy storage, and optoelectronic devices) as well as their excellent physical properties (such as electrical, mechanical, thermal, and optical properties) different from those of their bulk counterparts. The results show that the outputs of all optimization methods converge to ideal values with sufficiently long iterations and at different trials. However, when we consider the results of the statistical analyses of different trials under similar conditions, we observe that the method with the lowest error rate is the CMA-ES.

Keywords: Particle swarm optimization, Covariance matrix adaptation evolution strategies, NSGA-III, Two-dimensional transition metal dichalcogenides, Stillinger-Weber potential

1. INTRODUCTION

Transition metal dichalcogenides (TMDs) are layered materials having the chemical formula MX_2, where M is a transition metal atom (i.e., Mo, W, Nb) and X is a chalcogen atom (i.e., S, Se, Te). The TMDs can exhibit different electronic properties such as semiconducting (as in MoS_2 and WS_2), metallic (as in WTe_2 and NbS_2) and even superconducting (as in $NbSe_2$ and TaS_2) [1]. Nowadays with the advance of fabrication technologies, the low-dimensional (LD) structures of these materials can be fabricated and used in device and sensor applications. It has already proven that these materials have great potential to be used in various future technological applications such as transistors, high speed electronics, emitters, detectors, next-generation solar cells, LEDs, photodetectors, fuel cells, photocatalytic, flexible devices and touchscreen display panels that have more powerful, faster, smaller and more efficient characteristics than those currently existing [2-9].

Although there are many challenges in production, storage and conversion of the sustainable energy resources, the high-efficiency energy storage (e.g. supercapacitors and lithium ions batteries) and conservation (e.g. solar cells and electrocatalysis) initiatives are on the rise. It can be argued that both the market share and desire for working in the field will increase in the coming years [10] [11]. Moreover, the printed photodetectors in which such 2D nanomaterials is used, which is already in steady increase, estimated to reach a market share of 160 million dollars by 2023 [http://www.idtechex.com/research/articles/flexible-sensors-come-to-market-thanks-to-tens-of-millions-investment-00008464.asp]. Therefore, studying the physical properties of these TMD

monolayer structures is crucial to pave the way for the application of these materials in a wide variety of energy harvesting and storage related future technological applications [12], [2], [13].

For the scientific research of materials, computer simulation and modeling are indispensable methods in parallel with experimental work. First-principles calculation within the framework of density functional theory (DFT) has been extensively used in recent years to determine the properties of materials, and the obtained results have achieved a satisfactory level of accuracy that is comparable to experimental results [14], [15]. However, DFT simulations cannot be applied to large scale systems which is required to characterize size dependent material properties such as lattice thermal transport properties [16]. For this reason, molecular dynamics (MD) simulations are preferred in order to be able to model structures in a larger context. In order to realize a correct molecular dynamic simulation, it is necessary to predict a proper empirical potential to determine the interaction between the atoms (forces acting on the atoms). Developing the potential in accordance with the natural structure (metal, semiconductor, oxide, etc.) and purpose of the materials is crucial in determining the physical properties of the systems consistently. Many different potentials of interaction have been developed up to today and the obtained results are generally consistent with experimental data on different physical properties such as phase transition, structural, mechanical, and thermal properties of the materials studied [17]. These interactions are in general studied using conjugate gradient methods [18], and more recently, genetic algorithms [19], and particle swarm optimization [20], however, the results obtained using these methods are not generalized to the 2D material families and the accuracy of the generated potential on the vibrational properties of the crystal is limited. However, in order to understand the thermal dependent properties of a material (to determine its properties such as thermal expansion and conductivity) in a good way, it is very useful and intriguing to know its vibration properties, which mainly depends on the space derivative of the chosen potential. Therefore, it is of great importance to obtain the vibrational spectrum (phonon dispersion) of the system accurately with empirical potential, in addition to the basic properties such as lattice parameters, bond distances, elastic constants, and energy. In this optimization work, which we aim to develop empirical potential for some 2D TMDs, the target parameters consist of the equilibrium lattice parameter, bond distance and the phonon dispersion values, which we select for the specific wave vector points, that obtained by the first-principles calculation method.

The overall objective of this study is the systematical investigation of the performances of two real-valued metaheuristics, namely the particle swarm optimization and the covariance matrix adaptation evolution strategies and two different types of objective functions for simultaneous optimization of several different crystal characteristics. We implemented a methodical approach to determine the optimal parameter settings for the optimization of the structure of TMDs that is representative of two-dimensional materials that can be used for the future technological applications mentioned above. Moreover, we also studied the characterization of the 2D materials problem as a multiobjective optimization problem and for this, we used one of the state-of-the-art optimization approaches, nondominated sorting genetic algorithm, NSGA-III. For the purposes of our optimization procedure, we chose Stillinger-Weber type interatomic potential which is described in a simpler form and with fewer empirical parameters than the other interatomic potentials, such as Tersoff potential [21].

2. MODEL

The Stillinger-Weber (SW) potential is an empirical inter-atomic potential related to the distance and bond angles between atoms, embodying the two- and three-body interaction terms. At first, although this potential was developed for pure Si [23], afterwards it has been adapted to many other elements and compounds [24]. The form of the total energy of the crystal structure as a combination of two- and three-body interactions of the SW potential can be expressed as follows:

$$E = \sum_{i}\sum_{i>j}\phi_2\left(r_{ij}\right) + \sum_{i}\sum_{i\neq j}\sum_{k>j}\phi_3\left(r_{ij}, r_{ik}, \theta_{ijk}\right) \tag{1}$$

$$\phi_2\left(r_{ij}\right) = A_{ij}\left(\frac{B_{ij}}{r_{ij}^4} - 1\right)\exp\left[\frac{\rho_{ij}}{r_{ij} - r_{ij}^{\max}}\right] \tag{2}$$

$$\phi_3\left(r_{ij}, r_{ik}, \theta_{ijk}\right) = K\exp\left[\frac{\rho_{ij}}{r_{ij} - r_{ij}^{\max}} + \frac{\rho_{ik}}{r_{ik} - r_{ik}^{\max}}\right]\left(\cos\theta_{ijk} - \cos\theta_{0,ijk}\right)^2 \tag{3}$$

where, while ϕ_2 describe two-body potential term associated with scattering effect, ϕ_3 define the three-body potential term taking into account the bond angle effect. The summation indices j and k refer to the neighbor atoms within the maximum distance of r^{max} from atom i at center. The terms r_{ij} and r_{ik} are the bond length between the atom pairs i; j and i; k, respectively. And as the last term θ_{ijk} is the angle between the radial vectors from the central atom i towards neighbor atoms j and k ($\theta_{0,ijk}$ is the equilibrium angle). The pair dependent parameters A, B, ρ and K are fitting parameters to be determined by varying for two-body (bond stretch) and three-body (bond bend) energies.

Equation (1) represents the potential to be fitted, equations (2) and (3) are two-body and three body SW equations respectively. In our optimization process, we try to find the optimal values of the decision variables A, B, ρ, K, r_{ij}^{max}, r_{ik}^{max}, and $\theta_{0,ijk}$ to obtain the desired crystal characteristics.

In Figure 1, as an example of a single layer TMDs (MX_2, M = Mo, W, and X = S, Se), we schematically present the top and side view of the MX_2 configuration. In terms of location W and Mo are identical represented in red, and similarly, S and Se are identical for other materials, represented in green (MoS_2, $MoSe_2$, WS_2 and WSe_2). To describe this structure within the framework of the SW inter-atomic potential mentioned above, we need three types of stretching terms, namely for M-M, M-X, and X-X, where M and X are the corresponding transition metal and chalcogen atoms, respectively. Unlike the two-body interaction term, for the angle bending, it is essential to discriminate the chalcogen atoms due to the asymmetric chalcogen polyhedra around M (transition metal). The subscripts d and u are introduced to symbolize the chalcogen atoms below and above the M layer (M layer in Figure 1), respectively. Hereby, we define three types of three-body terms. The term $X_{d,(u)}$-M-$X_{d,(u)}$ is the angle between two chalcogens (both below or above the M-plane) and M atom at the center. The term M-$X_{d,(u)}$-M describes the angle between two M and one chalcogen (either below or above the M-plane) at the center. The term X_d-M-X_u is the angle between two chalcogens (one below and one above the M-plane) and M atom which is at the center. These angles that are labeled as θ_1, θ_2, and θ_3 are shown in Figure 1, respectively.
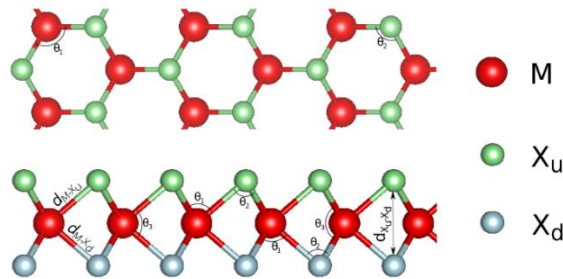
Figure 1. Generic structure of a single layer TMDs, top view (above) and side view (below)

3. OPTIMIZATION APPROACHES

In this study, we used two different evolutionary algorithms, one based on swarm intelligence (PSO) and the other is CMA-ES. Before going into implementation details of the algorithms, it is necessary to elaborate on the search space. As explained in detail in Section 2, we have a total of 12 decision variables, 9 of which to be used to calculate two-body SW equations, whereas the remaining decision

variables are needed to compute three-body SW equations. Throughout the experimentation procedure, we preferred to maintain a search space that is as wide as possible. Hence, we determined the lower and upper bounds on decision variables as summarized in Table 1 and Table 2.

Table 1. Lower (L) and upper (U) bounds (B) for two-body SW parameters in GULP format

Material / Interaction	Limit	A (eV) M-X	ρ (Å) M-X	B (Å⁴) M-X	A (eV) X-X	ρ (Å) X-X	B (Å⁴) X-X	A (eV) M-M	ρ (Å) M-M	B (Å⁴) M-M
MoS$_2$, MoSe$_2$, WS$_2$, WSe$_2$	LB	1.0	0.1	1.0	0.1	0.1	1.0	0.1	0.1	1.0
	UB	30.0	3.0	30.0	5.0	1.0	20.0	5.0	1.0	100.0

Table 2. Lower and upper bounds for three-body SW parameters in GULP format

Material / Interaction	Limit	K (eV) M-X$_{d,(u)}$-X$_{d,(u)}$; X$_{d,(u)}$-M-M	ρ_{12}, ρ_{13} (Å) M-X$_{d,(u)}$-X$_{d,(u)}$; X$_{d,(u)}$-M-M; M-X$_d$-X$_u$	K (eV) M-X$_d$-X$_u$
MoS$_2$, MoSe$_2$, WS$_2$, WSe$_2$	LB	0.1	0.1	0.1
	UB	20.0	2.0	50.0

The desired values of the variable parameters have been obtained by calculating the first-principles pseudopotential plane-wave solutions based on density functional theory and density functional perturbation theory using the VASP packet [25] [26] [27]. In order to minimize the periodic interactions caused by periodic boundary conditions for single layer structures, it has been assumed that these structures are placed in a supercell with a vacuum spacing of 20 Å along the z-direction. The calculation parameters have been determined by performing detailed test calculations to ensure a good convergence of total energy and force calculations. In this context, a plane wave basis set with 500 eV kinetic energy cut-off has been used for all monolayer structures. The Γ point centered 26×26×1 k-point mesh within the Monkhorst–Pack scheme has been used for the Brillouin zone integrations of the primitive cell. For the vibration frequencies, firstly, the force constants have been computed from the density function perturbation theory [28] by means of VASP, then the frequencies through these force constants have been obtained by using an open source code PHONOPY [29]. In addition, for all crystal structures, the phonon calculations have been carried out on 4×4×1 conventional supercell structure with Γ point centered 8×8×1 k-points grids for the Brillouin zone sampling.

3.1. PARTICLE SWARM OPTIMIZATION – PSO

The optimization techniques that emerged in the 1990s began to be inspired by the social interactions that are common in nature as well as the evolutionary process. One of these methods is Particle Swarm Optimization (PSO). PSO is inspired by the collective behavior observed in animal feed (fish and bird flocks [30]). Especially with the bird flocks, when a food or water supply are discovered by a member of a flock, the entire flock is known for its ability to turn to the corresponding source quite quickly. In the same way, when a fish flock meets a threat, all of the flock can quickly move away from the area where the threat exists. It is often seen that opportunities and threats are rapidly transmitted to all members of the herd in all animal species that act as a community. Kennedy and Eberhart have shown that these behaviors observed in nature can be an effective method of optimization [30]. In PSO, each particle is represented by three vectors (**x**, **p**, and **v**), and two objective functions ($f(\mathbf{x})$ and $f(\mathbf{p})$). In PSO, **x** vector represents the current position of the particle, **p** is the particle's best position, and **v** is the speed and motion direction of the particle. Lengths of these vectors are the same as each other, and each element in the vector corresponds to a decision variable. $f(\mathbf{x})$ represents the value of the objective function corresponding to the vector **x**, and $f(\mathbf{p})$ represents the objective function corresponding to the vector **p**. Each particle in the swarm uses its information from its search history and collectively gather the knowledge of the swarm to direct its own search, bringing it together in a vector **v** that determines the direction and speed of movement. A pseudocode of the PSO algorithm is provided below:

```
Algorithm1: PSO
```
```
for each particle i = 1, ..., S do
    Initialize the particles: xᵢ ~ U(bₗₒ, bᵤₚ)
    Initialize the particle's best known position: pᵢ ← xᵢ
    if f(pᵢ) < f(g) then
     Update the swarm's best known  position: g ← pᵢ
     Initialize the particle's velocity: vᵢ ~ U(-|bᵤₚ-bₗₒ|, |bᵤₚ-bₗₒ|)
end for
while a termination criterion is not met do:
    for each particle i = 1, ..., S do
       for each dimension d = 1, ..., n do
          Pick random numbers: rₚ, r_g ~ U(0,1)
          vᵢ,d ← K(vᵢ,d + φₚ rₚ (pᵢ,d-xᵢ,d) + φ_g r_g (g_d-xᵢ,d))
          xᵢ ← xᵢ + vᵢ
          if f(xᵢ) < f(pᵢ) then
            Update the particle's best known position: pᵢ ← xᵢ
          if f(pᵢ) < f(g) then
             Update the swarm's best known position: g ← pᵢ
       end for
    end for
end while
```

Figure 2. PSO Pseudocode

In the above pseudocode i represents the ith particle of the swarm, φ_1 and φ_2 are the coefficients governing the cognitive and social aspects of a particle's search, g represents the particle with the best objective function in the swarm, and d represents the dth dimension of the corresponding particle. The representation of U (0, 1) is a random number from a continuous uniform distribution with a lower bound of 0 and an upper bound of 1. In the PSO algorithm, the constriction coefficient κ is a function of φ_1 and φ_2 [31] and it is mainly used to provide control and stability on the growth of vector **v**. In the PSO algorithm, the current position (**x**) of the particle is changed at each iteration using the updated **v**, taking into account the social and cognitive effects. If the new position of the particle is better than the previously discovered ones by that particle, **p** and $f(\mathbf{p})$ are also updated. In this way the PSO continues its iterations until all the particles in the swarm (hopefully) converges to the same solution. The termination criterion generally depends on the problem at hand, but in general either a constant number of iterations or number of iterations that did not change the best known solutions are preferred.

As mentioned earlier, one of the cornerstones of the PSO algorithm is the communication between the solution candidates in the swarm. Different neighborhood definitions have been developed for this. The most commonly used neighborhood definitions are the ring topology and the global topology. In the ring topology, each particle can communicate with the other two particles closest to it, and each particle in the global topology can communicate with every other member of the swarm. Thus, when global topology is used, the swarm can be aware of the other members that discovers "good" solution alternatives more quickly. However, this may lead to premature convergence [31], which would prevent the PSO algorithm from achieving a better solution than expected.

3.2. COVARIANCE MATRIX ADAPTATION EVOLUTION STRATEGIES – CMA-ES

Evolutionary strategies (ES) belong to the class of evolutionary algorithms in nature-inspired search and optimization methods. Like other optimization methods in this class, it aims to achieve better results step by step by using mechanisms based on the theory of evolution such as recombination, mutation, and selection in a population of candidate solutions. The first studies on ES were carried out by Bienert, Rechenberg, and Schwefel, as stated in [32]. ES can be implemented in optimization problems where

decision variables are either continuous or discrete. ES can also be used in situations where there are constraints that limit solution space as well as in combinatorial optimization problems. These elasticities about the structure of decision variables are also valid for the objective function. The objective function can be defined for the ES by a mathematical function, an empirical model, or even physical measurements.

(μ, λ) and (μ + λ) are the most commonly used representations of an ES algorithm. In this notation, μ denotes the number of solution candidates in one generation of ES population, and λ denotes the number of new solutions to be derived using recombination and mutation from a generation. At each step of any ES algorithm, 6 to 8 times the number of solution candidates in a generation is generally derived as a new candidate solution [33]. In order to keep the population size constant, in (μ, λ) ES the new generation of solutions are selected among the λ newly generated candidate solutions, the next generation in (μ + λ) ES all the new solution candidates produced, as well as μ member of the previous generation struggle to survive to the next generation. For this reason, (μ + λ) -ES is called an "elitist" evolution strategy.

There are many different implementations of ES proposed in the literature. In this study, we prefer to use covariance matrix adaptation evolutionary strategy (CMA-ES), first introduced by [34], [35] as it is one of the most successful variants of the ES [32]. In CMA-ES, recombination is performed via weighted summation of the best μ individuals and the mutation is performed using a multivariate normal distribution. At each iteration, mutation parameters are updated according to the equations provided within the pseudocode. As with the most evolutionary search algorithms, the search is terminated when the maximum number function evaluations is reached. Pseudocode of the CMA-ES algorithm is provided below Bäck et al. [32]:

Algorithm2: CMA-ES

Initialization

Initialize $\langle \mathbf{X} \rangle$

$\mathbf{C} \leftarrow \mathbf{I}$

$\mathbf{p}_c, \mathbf{p}_\sigma \leftarrow \mathbf{0}$

$t \leftarrow 0$

while stopping criterion is not met **do**

    //sample and evaluate offspring

    $t \leftarrow t + 1$

    $\mathbf{B}$ and $\mathbf{D} \leftarrow$ eigendecomposition of $\mathbf{C}$

    for $i \in \{1, \ldots, \lambda\}$ **do**

        $\mathbf{z}_i \leftarrow N(\mathbf{0}, \mathbf{I})$

        $\mathbf{y}_i \leftarrow \mathbf{BDz}_i$

        $\mathbf{x}_i \leftarrow \langle \mathbf{x} \rangle + \sigma \mathbf{y}_k$

        $f_i \leftarrow f(\mathbf{x}_i)$

    end

    sort $\{\mathbf{x}_i\}$ with respect to $f(\mathbf{x}_i)$

    //internal update (path and parameters)

    $\langle \mathbf{y} \rangle \leftarrow \sum_{i=1}^{\mu} w_i y_{i:\lambda}$

    $\langle \mathbf{x} \rangle \leftarrow \langle \mathbf{x} \rangle + \sigma \langle \mathbf{y} \rangle = \sum_{i=1}^{\mu} w_i x_{i:\lambda}$

    $\mathbf{p}_\sigma \leftarrow (1 - c_\sigma)\mathbf{p}_\sigma + \sqrt{c_\sigma(2 - c_\sigma)\mu_{eff}} \mathbf{BD}^{-1}\mathbf{B}^T \langle \mathbf{y} \rangle$

    $\mathbf{p}_c \leftarrow (1 - c_c)\mathbf{p}_c + \sqrt{c_c(2 - c_c)\mu_{eff}} \langle \mathbf{y} \rangle$

    $\mathbf{C} \leftarrow (1 - c_1 - c_\mu)\mathbf{C} + c_1 \mathbf{p}_c \mathbf{p}_c^T + c_\mu \sum_{i=1}^{\mu} w_i \mathbf{y}_{i:\lambda} \mathbf{y}_{i:\lambda}^T$

    $\sigma \leftarrow \sigma \exp\left(\dfrac{c_\sigma}{d_\sigma}\left(\dfrac{\|p_\sigma\|}{E\|N(\mathbf{0}, \mathbf{I})\|} - 1 \right) \right)$

end

Figure 3. CMA-ES Pseudocode

In the pseudocode above, **x** represents the candidate solutions. $\mathbf{p}_c$ and $\mathbf{p}_\sigma$ represent the two evolution paths governing the covariance matrix adaptation and the global step size adaptation respectively, whereas **C** represents the covariance matrix. To improve readability, the mutation operation used in generating new candidate solutions is denoted by **y**. For the purposes of our implementation, of the CMA-ES is performed using the parameter settings as proposed by [36]. That is to say, the initial values of $\mathbf{C} \leftarrow \mathbf{I}$, $\mathbf{p}_c, \mathbf{p}_\sigma \leftarrow \mathbf{0}$ are used and the initial population is generated randomly between the lower and the upper bounds of the decision variables. At each iteration, $\lambda = 10\left(4 + \lfloor 3 + \ln(n) \rfloor\right)$ (where n represents the number of decision variables) new individuals are generated and $\mu = \lambda/2$ individuals are allowed in the next generation of candidate solutions. Lastly, in order to provide a fair comparison with PSO implementation in terms of total number of function evaluations, number of iterations for the CMA-ES is the same as of PSO implementation, which is 200.

3.3. OBJECTIVE FUNCTION

For the problem at hand, an "accurate" potential parameter set can be defined as the accuracy on the calculated physical properties of a crystal such as lattice constant, bond distance and phonon dispersion. It is possible to calculate these parameters using different approaches. One of these approaches is called first principles method, as implemented in VASP, The Vienna *a b initio* Simulation Package, [37] that gives results comparable with experimental measurements for a crystal systems having up to 500 atoms. Another method is known as molecular dynamics simulations, as implemented in a computer simulation package called GULP, General Utility Lattice Program, [38] that can be used to calculate temperature dependent physical properties of a crystal system having up to millions of atoms. However, the accuracy of Classical Molecular Dynamics simulations significantly depends on the empirical potential parameter set used. Therefore, in order to produce the required parameter set which can be used for large scale material simulations, we chose a crystal system which can be simulated using first principles approach, in other words, VASP. Then, we predicted the parameter set that provides us with the results having minimal deviation from the VASP calculations by the fitting algorithms mentioned above. Here, in order to perform MD simulations GULP program is employed. Thus, the objective function of the problem is defined as:

$$f(\mathbf{x}) = \sum_{j=1}^{J} w_j \frac{\left| t_j - a_j \right|}{t_j} \tag{4}$$

Where $f(\mathbf{x})$ is the objective function, t_j is the value of the jth crystal characteristic obtained through VASP, a_j is the value of the jth crystal characteristic obtained through GULP and w_j is the weight of the jth crystal characteristic in the objective function. The first characteristic defined in the objective function is the bond distance, and the second one is the lattice constant. The remaining 28 characteristics define the phonon dispersion frequencies over the high-symmetry directions of the Brillouin zone (minimum unsymmetrical crystal Fourier space in which lattice vibration waves can be defined). Phonon dispersion diagrams are especially important in assessing the high temperature crystal properties such as thermal conductivity. Since it is not possible to represent the dispersion diagrams in the objective function as a whole, we choose to sample some critical points representative of the phonon diagrams. The weights are determined by considering the importance of different crystal characteristics relative to each other. Throughout the experimentation phase, two different version of the objective function are used. In the first version, all crystal characteristics are treated as equal, which makes $w_j = 1$ for all j. In the second version, the characteristics considered to be the least important are assigned $w_j = 1$, moderately important characteristics has a weight value of 2, and the most important characteristics has a weight value of 4. In our calculations, the highest weights are assigned to the optic phonon frequencies, along with the bond distance and lattice constant. During experimentation, we attempted to answer two

questions: whether there is a statistically significant difference between the two optimization procedures in terms of performance, and whether the use of weights contribute significantly to the performance of these optimization algorithms. Answers to both these questions are discussed in detail in the next section.

3.4. STATISTICAL ANALYSIS

In this section, we provide the results of the experimentation performed using two different optimization procedures, namely the PSO and the CMA-ES, and two different objective functions, one with weights, and the other without weights. Thus, we obtained four different experimental settings. For each experimental setting, we ran the algorithm 30 times and the best solution obtained at each trial are recorded. We then compared the performance of each configuration by evaluating the average, best (minimum) and worst (maximum) objective function values as summarized in Table 3. Both PSO and CMA-ES are coded using MATLAB and ran on a MacBook with a 1.7 GHz Intel Core i7 processor with 8 GB of RAM.

In Table 3 below, we summarize the results for four different single layer crystal structures, namely MoS_2, $MoSe_2$, WS_2 and WSe_2 obtained using PSO and CMA-ES and weighted and unweighted versions of the objective function. In order to make a fair comparison between the weighted and the unweighted versions of the objective functions, the results obtained using the weighted version of the objective function is converted into unweighted version by removing the effect of weights from the objective function values. It is observed that, the sum of deviations from the desired crystal characteristics, on average, is smaller when the CMA-ES is used as the optimization procedure, except for the MoS_2. It can be also verified from Table 3 that the range between the minimum and the maximum values of the objective functions are much smaller with the CMA-ES.

We also provide the box plots for the ranges of best solutions identified in different trials of each optimization method using two different versions of the objective functions in Figure 4 and Figure 5. As can be verified from the boxplots, the range of the "best" solutions identified by the PSO algorithm is much greater than the ones identified by the CMA-ES. The difference between the maximum and the minimum objective function values is even more pronounced for the PSO when the weighted objective function is used, however, the effect of weighted objective function on CMA-ES is hardly detectable. It is our conjecture that this difference can be attributable to the new candidate solution generation mechanisms of PSO and CMA-ES. In PSO, new candidate solutions are generated by using the information on the particle's best position and the best position of the overall swarm using uniform distribution, whereas in CMA-ES the new solutions are generated using a multigaussian distribution. Consequently, once CMA-ES identifies a promising region in the objective function space, this information quickly propagates across the whole population, whereas for PSO this would take more time and number of iterations. On one hand, this helps PSO to find better solutions occasionally, and on the other hand, it also leads to keep the search algorithm in exploration stage too long, as indicated by the greater range between maximum and minimum values of the identified "best" solutions at different trials of the PSO algorithm.

Table 3. Summary of results for MoS_2, $Mo Se_2$, WS_2, WSe_2

Objective Function		MoS₂		MoSe₂		WS₂		WSe₂	
		PSO	CMA-ES	PSO	CMA-ES	PSO	CMA-ES	PSO	CMA-ES
$w_j = 1$	Max (Worst)	1.118	0.843	1.924	1.676	2.842	0.811	1.584	0.882
	Min (Best)	0.747	0.784	1.489	1.570	0.711	0.743	0.781	0.814
	Average	0.800	0.809	1.679	1.622	0.879	0.778	0.931	0.848
$w_j \neq 1$	Max (Worst)	2.346	0.843	2.276	1.676	1.291	0.873	1.859	0.955
	Min (Best)	0.763	0.787	1.562	1.577	0.745	0.732	0.808	0.816
	Average	0.847	0.808	1.673	1.625	0.852	0.789	1.005	0.881
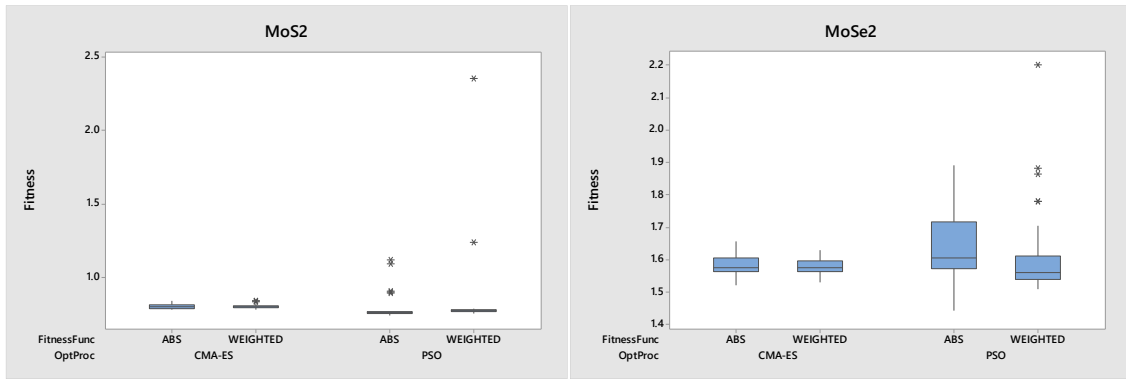
Figure 4. Boxplots of fitness functions, MoS_2 and $MoSe_2$ - CMA-ES vs PSO
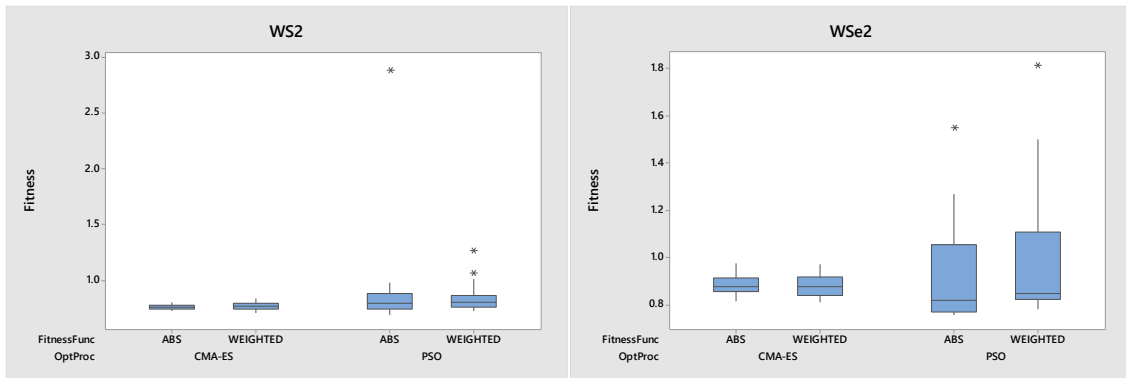
Figure 5. Boxplots of fitness functions, WS_2 and WSe_2 - CMA-ES vs PSO

In order to assess whether there is a significant difference between the two optimization procedure and, the two objective functions, ANOVA test is performed. Thus, there are two factors with two levels each e.g. optimization procedure (PSO vs CMA-ES) and the type of the objective function (unweighted vs weighted). All statistical analyses are performed using 95% level of confidence. We also considered the interaction between the optimization procedure and the type of the objective function. For brevity, we summarized the results as in Table 4, only reporting the computed F statistics and the corresponding *p* value (in parentheses). We also run ANOVA tests to see specifically if there is a significant contribution of the factors defined above on Lattice Constant, Bond Distance and Phonon Frequencies as reported in Table 5, Table 6 and Table 7 respectively.

Overall, CMA-ES outperforms PSO in three out of four crystal structures, namely, $MoSe_2$, WS_2, and WSe_2, as observed in Table 4. Also can be verified from Table 4 that the type of the objective function has no statistically significant effect on the performance of the algorithms on overall level. In order to continue with the statistical analyses on further detail, we decompose the overall objective function into its constituents as Lattice Constant, Bond Distance and Phonon Frequencies. The weight value used for lattice constant and bond distance is $w_1 = w_2 = 4$. For the lattice constant, the use of weights proved to be an effective tool in reducing the average deviation from the corresponding ideal value of the lattice constant for both MoS_2 and $MoSe_2$ whereas for the other two crystals the use of weights did not contribute significantly to the same purpose (Table 5). For the lattice constant, the only significant difference between optimization approaches found for WS_2, which is in favor of CMA-ES again. The effect of the use of weights to reduce the average deviation from the ideal values are even more pronounced in the case of bond distance, as can be verified from Table 6. The objective function with the weights produced a smaller deviation from the ideal values of all bond distances for all crystal structures considered. For the bond distance alone, there is no statistically significant difference between

the performances of CMA-ES and PSO for all crystals except for $MoSe_2$. Lastly, for the sum of the deviations from the ideal phonon frequency values, MoS_2 and $MoSe_2$, favors the weighted objective function, whereas for the other two crystals there is no statistically significant difference present in the type of objective function Table 7. For the frequencies again, CMA-ES comes forward as the choice of optimization procedure for $MoSe_2$, as summarized in Table 7.

When all these results are considered as a whole, the CMA-ES with objective function in which $w_j = 1$ for all characteristics $j = 1, ..., 30$ turns out to be the best option for the optimization of the crystal structures considered in this study. The type of the objective function is only significant for the deviations from the ideal bond distance values across all crystals, however, even in the worst case the actual deviation from the ideal value of the bond distance is only around 5%, which is well within the acceptable range. For the sake of simplicity and the greater potential of convergence to a solution that is acceptable from the view of all crystal characteristics, CMA-ES with unweighted objective function turns out to be a better option.

Table 4. Summary of results for MoS_2, $MoSe_2$, WS_2, WSe_2

Overall	PSO vs CMA-ES (1)	Type of Obj. Func. (2)	(1)×(2)
MoS_2	0.30 (0.59)	0.69 (0.41)	0.71 (0.40)
$MoSe_2$	**8.77 (0.00)**	0.00 (0.94)	0.05 (0.83)
WS_2	**5.20 (0.02)**	0.05 (0.82)	0.29 (0.59)
WSe_2	**11.12 (0.00)**	2.94 (0.09)	0.01 (0.50)

Table 5. Summary of results for MoS_2, $MoSe_2$, WS_2, WSe_2

Lattice Constant	PSO vs CMA-ES (1)	Type of Obj. Func. (2)	(1)×(2)
MoS_2	0.22 (0.64)	**4.98 (0.03)**	1.00 (0.32)
$MoSe_2$	0.8 (0.38)	**13.65 (0.00)**	**5.75 (0.02)**
WS_2	3.80 (0.05)	3.61(0.06)	2.43 (0.12)
WSe_2	**13.75 (0.00)**	1.73 (0.19)	3.32 (0.07)

Table 6. Summary of results for MoS_2, $MoSe_2$, WS_2, WSe_2

Bond Distance	PSO vs CMA-ES (1)	Type of Obj. Func. (2)	(1)×(2)
MoS_2	**7.06 (0.01)**	**40.83 (0.00)**	1.68 (0.20)
$MoSe_2$	1.47 (0.23)	**62.60 (0.00)**	0.58 (0.45)
WS_2	**8.05 (0.00)**	**4.17 (0.04)**	**4.19 (0.04)**
WSe_2	**20.35 (0.00)**	**21.42 (0.00)**	0.64 (0.43)

Table 7. Summary of results for MoS_2, $MoSe_2$, WS_2, WSe_2

Frequencies	PSO vs CMA-ES (1)	Type of Obj. Func. (2)	(1)×(2)
MoS_2	0.00 (0.99)	**4.41 (0.04)**	1.22 (0.27)
$MoSe_2$	**5.86 (0.02)**	**6.00 (0.02)**	0.00 (0.97)
WS_2	2.97 (0.09)	0.07 (0.79)	0.00 (0.95)
WSe_2	3.49 (0.06)	3.40 (0.07)	1.87 (0.17)

We also provide phonon frequency diagrams for all four crystals using PSO and CMA-ES along with weighted and not weighted objective functions in Figure 6 for MoS_2, Figure 7 for $MoSe_2$, Figure 8 for WS_2, and Figure 9 for WSe_2. Overall, both methods and objective functions produce quite satisfactory results with respect to phonon dispersion diagrams, however, based on the statistical analyses results,

we can claim that the CMA-ES with not weighted objective function provides smaller deviations across all characteristics.
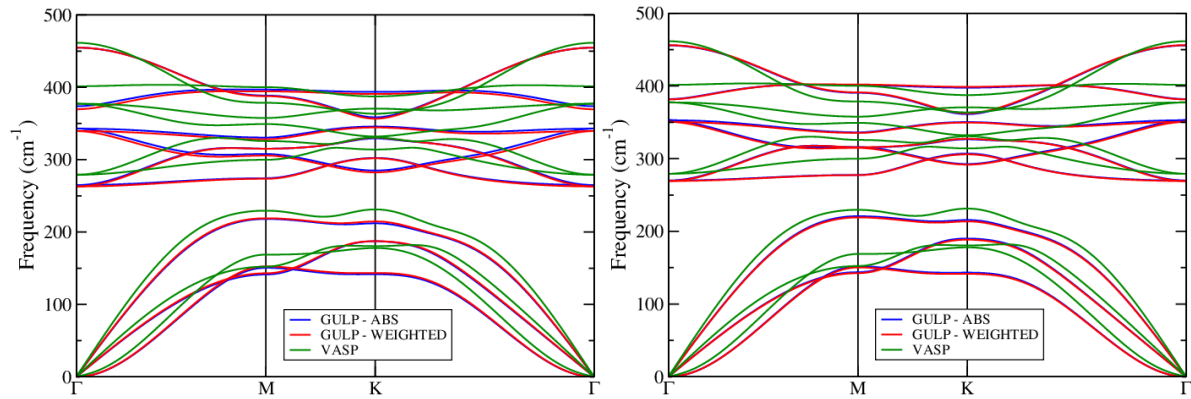

Figure 6. Phonon dispersions of single-layer MoS_2 structure along high-symmetry directions of the Brillouin zone: CMA-ES vs PSO (GULP results are for the SW parameters giving the best results belonging to the corresponding optimization method)
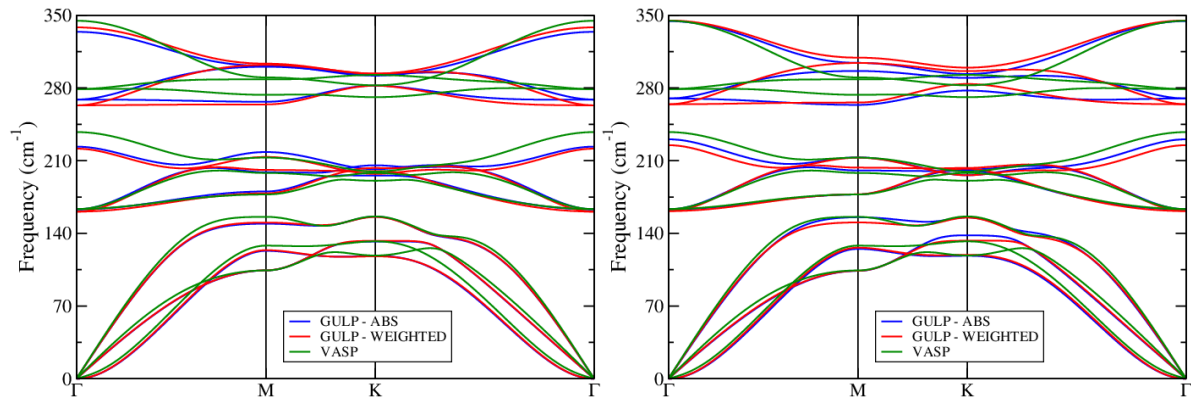
Figure 7. Phonon dispersions of single-layer $MoSe_2$ structure along high-symmetry directions of the Brillouin zone: CMA-ES vs PSO (GULP results are for the SW parameters giving the best results belonging to the corresponding optimization method)
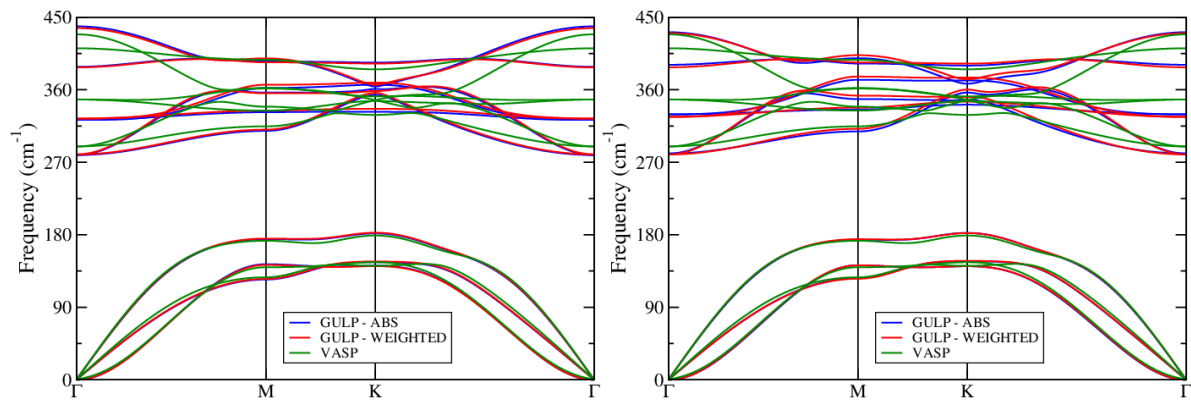
Figure 8. Phonon spectrums of monolayer WS_2 along the ΓMKΓ direction in the Brillouin zone: CMA-ES vs PSO (GULP results are for parameter set with the lowest error obtained using the relevant optimization method)
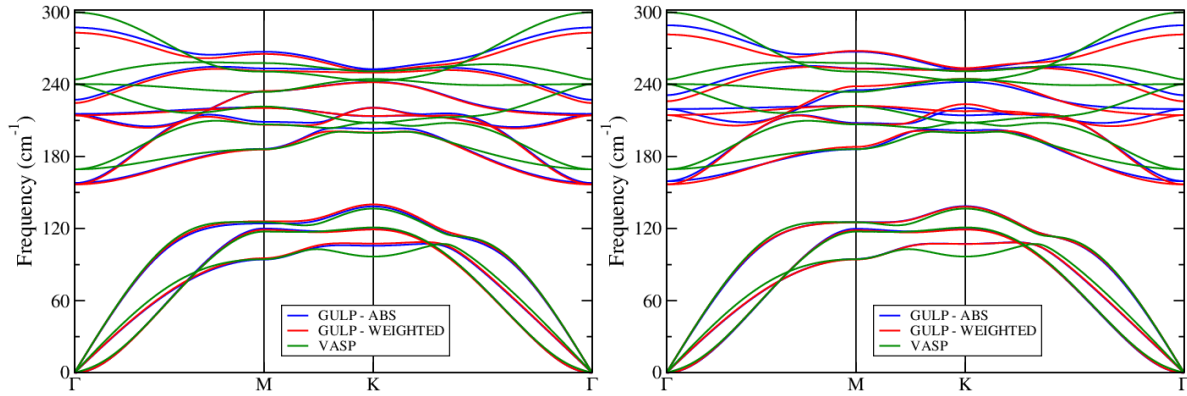
Figure 9. Phonon spectrums of monolayer WSe$_2$ along the ΓMKΓ direction in the Brillouin zone: CMA-ES vs PSO (GULP results are for parameter set with the lowest error obtained using the relevant optimization method).

4. MULTIOBJECTIVE MODEL

As explained previously, the objective function is composed of thirty different characteristics; bond distance, lattice constant, 28 points sampled from the phonon frequency diagrams that represent acoustic and optic frequencies. When there are many objectives to be considered simultaneously, multiobjective optimization techniques can also be used, in addition to the single objective approach discussed in the previous section. By doing so, we can also have an understanding of whether these objectives are conflicting or not. Another motivation for implementing a multiobjective optimization method is that using "Pareto optimal" solutions, the decision maker can assess the trade-offs to choose a solution among a set of alternative solutions to specifically minimize the deviation from the desired characteristic values and observe the worsening readily on the remaining objective functions. For this reason, we also created a three objective model and used NSGA-III as explained in [39] to solve the three-objective model. Although Deb and Jain stated that the NSGA-III is developed for optimization problems with four or more objective functions, they also presented test problems with three objective functions as well [39] with satisfactory results. A pseudocode of the NSGA-III algorithm is provided in Figure 10. As with most other multiobjective evolutionary algorithms (MOEA), NSGA-III also uses recombination and mutation to generate new offspring. NSGA-III specifically addresses two important issues: (1) NSGA-III uses a special domination rule based on adaptively discretizing the Pareto-optimal front to find a well-distributed set of points in the Pareto-optimal front and (2), instead of using predefined multiple search directions, multiple reference points are specified in NSGA-III beforehand to find a set of widely distributed Pareto-optimal points. In addition, recombination strategies are tuned to ensure creating near-parent solutions. For the details of the NSGA-III procedure, the reader is referred to [39].

For our problem, we defined three objective function. The first objective function is defined as the average normalized deviation of the bond distance and lattice constant values from their desired values. The second and third objective functions are the average normalized deviation of 16 optic and 12 acoustic phonon frequencies from their desired values, respectively. With these settings, NSGA-III is allowed to run for 5000 iterations with a population size of 100 individuals, creating 100 offspring at each iteration. Following the recommendations of Deb and Jain [39], we used simulated binary crossover, SBX, for creating offspring, and these offspring are mutated with a probability of 0.02. NSGA-III is allowed to run for a much lengthier period of time in comparison to PSO and CMA-ES, and the Pareto optimal solutions found as a result are then compared to the results reported in the previous section. The main reason for not having repetitions for the NSGA-III is the fact that all the solutions identified at the end of a single run are all optimal solutions in the Pareto-preference sense.

```
Algorithm3: NSGA-III procedure at generation t
```

Input: H structured reference points Z^s, parent population P_t

Output: P_{t+1}

```
Begin
```

$S_t \leftarrow \emptyset, i \leftarrow 1;$

$Q_t \leftarrow \texttt{Variation}\ (P_t);$

$R_t \leftarrow P_t \cup Q_t;$

$(F_1, F_2, ...) \leftarrow \texttt{Non-dominated\_Sort}\ (R_t);$

while $|S_t| \geq N$ **do**

$\qquad S_t \leftarrow S_t \cup F_i;\ i \leftarrow i+1;$

end

$F_l \leftarrow F_i;$ //Last front to be included

if $|S_t| = N$ then

$\qquad P_{t+1} \leftarrow S_t;$

else

$\qquad P_{t+1} \leftarrow \bigcup_{j=1}^{i-1} F_j$

$\qquad$ //Number of points to be chosen from F_l

$\qquad K \leftarrow N - |P_{t+1}|;$

$\qquad$ //Normalize objectives and create reference set Z^r

$\qquad \texttt{Normalize}\ (F^M;\ S_t;\ Z^r;\ Z^s);$

$\qquad$ //Associate each member s of S_t with a reference point

$\qquad$ // $\pi(s)$: closest reference point

$\qquad$ // $d(s)$: distance between s and $\pi(s)$

$\qquad [\pi(s), d(s)] \leftarrow \texttt{Associate}\ (S_t;\ Z^r);$

$\qquad$ //Compute niche count of reference point $j \in Z^r$

$\qquad \rho_j \leftarrow \sum_{s \in S_t / F_l} \big((\pi_s = j)?1:0\big);$

$\qquad$ //Choose K members one at a time from F_l to construct P_{t+1}

$\qquad \texttt{Niching}\ (K, \rho_j, \pi(s), d(s), Z^r, F_l, P_{t+1});$

end if

end

Figure 10. NSGA-III Pseudocode

4.1. STATISTICAL ANALYSIS: CMA-ES VS NSGA-III

The results obtained from NSGA-III is only compared with that of CMA-ES. We already established in the previous sections that CMA-ES produces better results compared to PSO. In addition, the unweighted objective function performs better than the weighted objective function, so the statistical analysis performed in this section only compares the results obtained using CMA-ES with unweighted objective function with that of NSGA-III.

We begin statistical analysis by comparing the maximum (worst) minimum (best) and average deviations from the desired values of crystal characteristics, as summarized in Table 8. As can be verified from the *t*-stat column of Table 8, overall the sum of deviations from the desired crystal characteristics on average, is smaller with the CMA-ES in comparison to NSGA-III, with 95% level of confidence. For WS_2, NSGA-III provides slightly better results than CMA-ES, however, the difference is not significant according to the *t*-test performed. We also provide the boxplots on overall results summarized in Table 9 on Figure 11 for MoS_2 and $MoSe_2$ and for WS_2 and WSe_2 on Figure 12.

Table 8. Comparison of Results: CMA-ES vs NSGA-III on the basis of defined objective functions for NSGA-III

		Overall		
	Optimization Procedure	CMA-ES	NSGA-III	*t* stat.
MoS_2	Maximum	0.843	2.097	
	Minimum	0.784	0.855	**-10.56 (0.000)**
	Mean	0.809	0.947	
$MoSe_2$	Maximum	1.676	1.832	
	Minimum	1.570	1.633	**-6.52 (0.000)**
	Mean	1.622	1.656	
WS_2	Maximum	0.811	0.810	
	Minimum	0.743	0.748	1.14(0.262)
	Mean	0.778	0.774	
WSe_2	Maximum	0.882	1.276	
	Minimum	0.814	0.997	**-33.52 (0.000)**
	Mean	0.848	1.045	
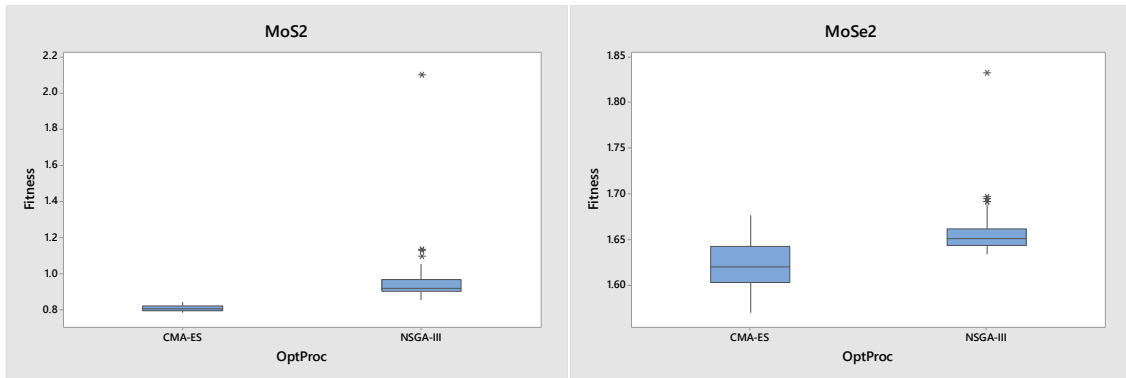
Figure 11. Boxplots of fitness functions, MoS_2 and $MoSe_2$ - CMA-ES vs NSGA-III
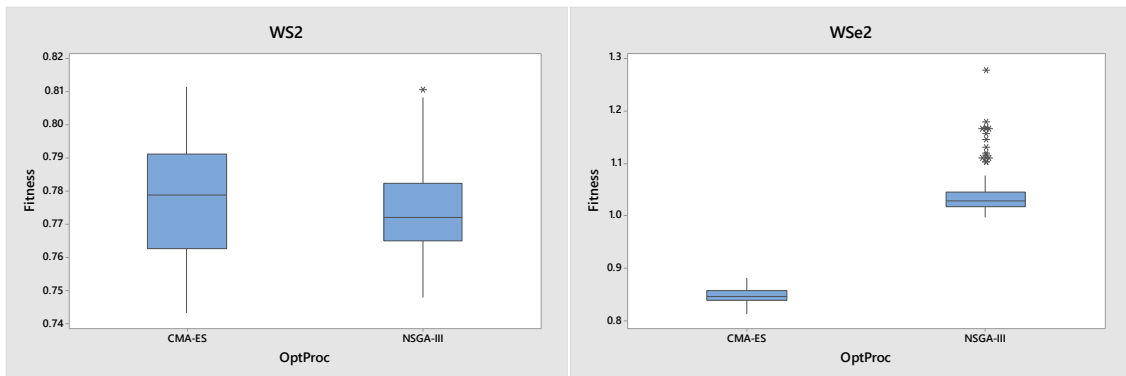
Figure 12. Boxplots of fitness functions, WS_2 and WSe_2 - CMA-ES vs NSGA-III

In addition to the overall analysis, we decomposed the crystal characteristics in two different ways. First, we converted the results obtained using CMA-ES to that of NSGA-III and listed below in Table 9. For the first objective function (average deviation from the bond distance and lattice constant), the average deviation from the desired values of those characteristics are smaller with NSGA-III compared to CMA-

ES, however, for the second and third objective functions (deviations from the optic and acoustic frequencies, respectively), CMA-ES provides better results compared to NSGA-III, in general. As can be further verified from Table 10, the difference for the first objective function is in favor of NSGA-III, whereas for the third objective function, the average deviation from the desired values of the frequencies are smaller with CMA-ES with 95% level of confidence across all crystal structures experimented with. For the second objective function, on the other hand, CMA-ES produced better results for MoS_2 and NSGA-III produced better results for $MoSe_2$. For the remaining two crystals, there is no statistically significant difference between the two.

Table 9. Comparison of Results: CMA-ES vs NSGA-III on the basis of defined objective functions for NSGA-III

Variable		1st Objective Function		2nd Objective Function		3rd Objective Function	
	Optimization Procedure	CMA-ES	NSGA-III	CMA-ES	NSGA-III	CMA-ES	NSGA-III
MoS_2	Maximum	0.050	0.025	0.036	0.104	0.024	0.048
	Minimum	0.008	0.003	0.028	0.032	0.017	0.022
	Mean	0.028	0.012	0.032	0.035	0.020	0.031
$MoSe_2$	Maximum	0.085	0.006	0.035	0.027	0.100	0.120
	Minimum	0.011	0.001	0.021	0.023	0.088	0.100
	Mean	0.037	0.003	0.028	0.025	0.092	0.104
WS_2	Maximum	0.055	0.012	0.036	0.035	0.023	0.025
	Minimum	0.009	0.001	0.026	0.031	0.016	0.016
	Mean	0.024	0.003	0.032	0.032	0.018	0.021
WSe_2	Maximum	0.055	0.011	0.032	0.046	0.031	0.062
	Minimum	0.015	0.001	0.025	0.026	0.022	0.043
	Mean	0.032	0.005	0.029	0.028	0.027	0.049

Table 10. Summary of t-tests for MoS_2, $MoSe_2$, WS_2, WSe_2: the difference between CMA-ES and NSGA-III

	Bond Distance + Lattice Constant	Optic Frequencies	Acoustic Frequencies
MoS_2	**7.53 (0.00)**	**-2.67 (0.01)**	**-19.29 (0.00)**
$MoSe_2$	**9.69 (0.00)**	**4.24 (0.00)**	**-22.39 (0.00)**
WS_2	**9.81 (0.00)**	-1.78 (0.09)	**-5.29 (0.00)**
WSe_2	**13.60 (0.00)**	1.50 (0.14)	**-38.70 (0.00)**

The second decomposition is performed based on the lattice constant, the bond distance and the phonon frequencies and summarized in Table 11. As can be verified both Table 11 and Table 12, NSGA-III performs better than the CMA-ES on the basis of both the bond distance and lattice constant, however, on the basis of average deviation from the average frequencies, CMA-ES is the better performing optimization method. Considering the fact that average frequency considers 28 crystal characteristics simultaneously, once again CMA-ES stands out as a better option.

Table 11. Comparison of Results: CMA-ES vs NSGA-III on the basis of Lattice Constant, Bond Distance and Average Phonon Frequencies

	Variable	Lattice Constant		Bond Distance		Average Frequency	
	Optimization Procedure	CMA-ES	NSGA-III	CMA-ES	NSGA-III	CMA-ES	NSGA-III
MoS_2	Maximum	0.029	0.009	0.088	0.041	0.028	0.075
	Minimum	0.000	0.000	0.001	0.000	0.025	0.030
	Mean	0.010	0.002	0.045	0.021	0.026	0.033
$MoSe_2$	Maximum	0.044	0.005	0.130	0.006	0.057	0.063
	Minimum	0.000	0.000	0.008	0.002	0.050	0.056
	Mean	0.013	0.002	0.062	0.004	0.054	0.057
WS_2	Maximum	0.055	0.012	0.082	0.013	0.027	0.028
	Minimum	0.001	0.000	0.002	0.001	0.024	0.025
	Mean	0.015	0.001	0.033	0.005	0.026	0.027
WSe_2	Maximum	0.065	0.007	0.115	0.015	0.032	0.044
	Minimum	0.000	0.000	0.002	0.000	0.025	0.034
	Mean	0.027	0.002	0.045	0.007	0.029	0.036

Table 12. Summary of t-tests for MoS2, MoSe2, WS2, WSe2: The difference between CMA-ES and NSGA-III

	Bond Distance	Lattice Constant	Frequencies
MoS_2	5.44 (0.00)	5.37 (0.00)	**-13.47 (0.00)**
$MoSe_2$	9.94 (0.00)	5.03 (0.00)	**-10.02 (0.00)**
WS_2	7.22 (0.00)	5.85 (0.00)	**-4.13 (0.00)**
WSe_2	9.25 (0.00)	6.53 (0.00)	**-18.30 (0.00)**

In Figure 13 and Figure 14, we provide examples of the phonon frequency diagrams for the crystals considered in this study. For the acoustic frequencies which is shown on the lower half of the diagrams, NSGA-III closely matches the results obtained from VASP. However, for the optic frequencies, the deviations are much higher in all four crystals.
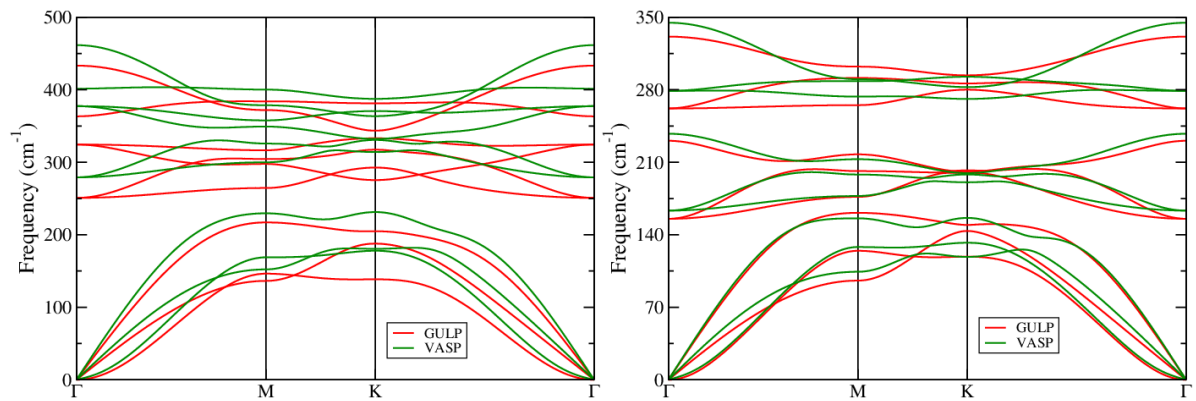
Figure 13. Phonon frequency diagrams drawn along high-symmetry directions in reciprocal space according to VASP and GULP (for parameter set that has the lowest error obtained using the NSGA-III method) results: MoS2 and MoSe2
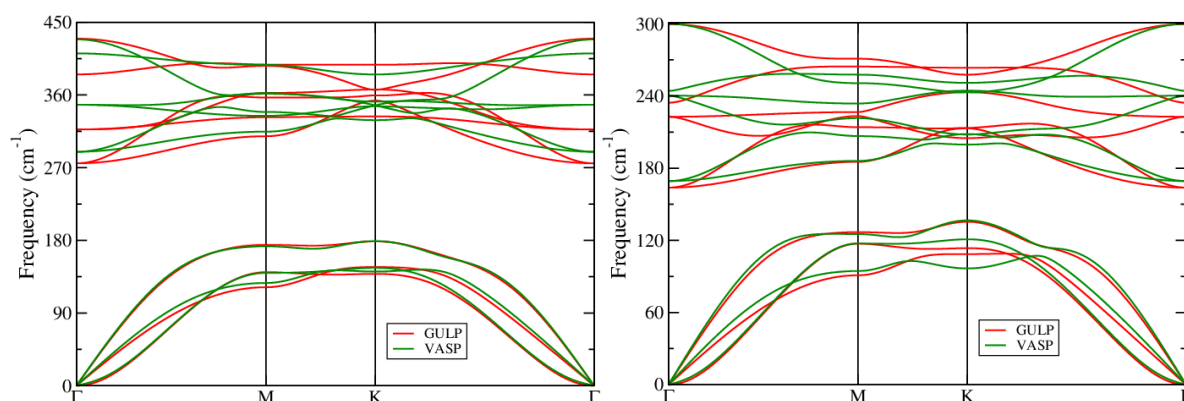
Figure 14. Phonon frequency diagrams drawn along high-symmetry directions in reciprocal space according to VASP and GULP (for parameter set that has the lowest error obtained using the NSGA-III method) results: WS$_2$ and WSe$_2$

5. CONCLUSION AND FUTURE WORK

In this study, we implemented three different evolutionary optimization methods to the theoretical optimization of two-dimensional transition metal dichalcogenides characteristics. To the best of our knowledge, this is the first study in the literature that systematically compares various optimization methods' performances on the problem. According to the results, CMA-ES turns out to be the best option among the alternatives in optimizing the theoretical characteristics of such materials. Researchers in the field can benefit from both results reported here, and the optimization approach as a whole, as the solution approach presented here is quite new for the field of theoretical/computational physics.

As stated previously, in implementing multiobjective optimization model, one of our motivations was to be able to obtain even smaller average deviation values in objective functions. However, even the smallest values for the defined objective functions for the optic and acoustic frequencies obtained using NSGA-III turned out to be greater than the ones obtained using CMA-ES. Even then, by carefully selecting and defining more objective functions, one can possibly achieve smaller deviations on a specific set of crystal characteristics, which is the case for the first objective function of the NSGA-III model. From that respect, we cannot claim that we fully exploited the potential of the NSGA-III algorithm.

ACKNOWLEDGEMENTS

REFERENCES

[1] Chhowalla M, et al., The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets. Nature Chemistry, 2013; 5: p. 263.

[2] Radisavljevic B, et al., Single-layer MoS$_2$ transistors. Nat Nanotechnol, 2011; 6(3): p. 147-50.

[3] Lopez-Sanchez O, et al., Ultrasensitive photodetectors based on monolayer MoS$_2$. Nat Nanotechnol, 2013; 8(7): p. 497-501.

[4] Cao T, et al., Valley-selective circular dichroism of monolayer molybdenum disulphide. Nature Communications, 2012; 3: p. 887.

[5] Jariwala D., et al., Emerging Device Applications for Semiconducting Two-Dimensional Transition Metal Dichalcogenides. ACS Nano, 2014; 8(2): p. 1102-1120.

[6]   Wang Q.H., et al., Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. Nature Nanotechnology, 2012; 7: p. 699.

[7]   Radisavljevic B., M.B. Whitwick, and A. Kis, Small-signal amplifier based on single-layer MoS2. Applied Physics Letters, 2012; 101(4): p. 043103.

[8]   Liu H. and P.D. Ye, MoS$_2$ Dual-Gate MOSFET With Atomic-Layer-Deposited Al2O3 as Top-Gate Dielectric. IEEE Electron Device Letters, 2012; 33(4): p. 546-548.

[9]   Fang H., et al., High-performance single layered WSe(2) p-FETs with chemically doped contacts. Nano Lett, 2012;12(7): p. 3788-92.

[10] Pumera M, Graphene-based nanomaterials for energy storage. Energy & Environmental Science, 2011;4(3): p. 668-674.

[11] Wang H, H. Feng, and J. Li, Graphene and Graphene-like Layered Transition Metal Dichalcogenides in Energy Conversion and Storage. Small, 2014; 10(11): p. 2165-2181.

[12] Huang J.-K, et al., Large-Area Synthesis of Highly Crystalline WSe$_2$ Monolayers and Device Applications. ACS Nano, 2014; 8(1): p. 923-930.

[13] Gutiérrez H.R., et al., Extraordinary Room-Temperature Photoluminescence in Triangular WS$_2$ Monolayers. Nano Letters, 2013; 13(8): p. 3447-3454.

[14] Hohenberg, P. and W. Kohn, Inhomogeneous Electron Gas. Physical Review, 1964. 136(3B): p. B864-B871.

[15] Aykol M. and C. Wolverton, Local environment dependent GGA+U method for accurate thermochemistry of transition metal compounds. Physical Review B, 2014; 90(11): p. 115105.

[16] Payam N. and J.S. David, Thermal conductivity of single-layer WSe$_2$ by a Stillinger–Weber potential. Nanotechnology, 2017; 28(7): p. 075708.

[17] Rapaport D.C., The Art of Molecular Dynamics Simulation. 2 ed. 2004; Cambridge: Cambridge University Press.

[18] Andrei N., Conjugate gradient Algorithms for Molecular Formation under pairwise Potential Minimization, in Proceedings of the Fifth Workshop on Mathematical Modelling of Environmental and Life Sciences Problems. 2006; Romania.

[19] Solomon J., et al., Method and advantages of genetic algorithms in parameterization of interatomic potentials: Metal oxides. Computational Materials Science, 2014; 81: p. 453-465.

[20] Voglis C., et al., A parallel hybrid optimization algorithm for fitting interatomic potentials. Applied Soft Computing, 2013; 13(12): p. 4481-4492.

[21] Tersoff J., New empirical model for the structural properties of silicon. Physical Review Letters, 1986; 56(6): p. 632-635.

[22] Liang T., S.R. Phillpot, and S.B. Sinnott, Parametrization of a reactive many-body potential for Mo--S systems. Physical Review B, 2009; 79(24): p. 245110.

[23] Stillinger F.H. and T.A. Weber, Computer simulation of local order in condensed phases of silicon. Physical Review B, 1985; 31(8): p. 5262-5271.

[24] Ichimura M., Stillinger-Weber potentials for III–V compound semiconductors and their application to the critical thickness calculation for InAs/GaAs. physica status solidi (a), 1996; 153(2): p. 431-437.

[25] Blöchl P.E., Projector augmented-wave method. Physical Review B, 1994; 50(24): p. 17953-17979.

[26] Kresse G. and J. Hafner, Ab initio molecular dynamics for open-shell transition metals. Physical Review B, 1993; 48(17): p. 13115-13118.

[27] Kresse G. and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. Physical Review B, 1996; 54(16): p. 11169-11186.

[28] Baroni S., et al., Phonons and related crystal properties from density-functional perturbation theory. Reviews of Modern Physics, 2001; 73(2): p. 515-562.

[29] Togo A., F. Oba, and I. Tanaka, First-principles calculations of the ferroelastic transition between rutile-type and $CaCl_2$ type SiO_2 at high pressures. Physical Review B, 2008. 78(13): p. 134106.

[30] Kennedy J.F., R.C. Eberhart, and Y. Shi, Swarm intelligence. The Morgan Kaufmann series in evolutionary computation. 2001, San Francisco: Morgan Kaufmann Publishers. xxvii, 512 p.

[31] Clerc M. and J. Kennedy, The particle swarm - explosion, stability, and convergence in a multidimensional complex space. Trans. Evol. Comp, 2002; 6(1): p. 58-73.

[32] Bäck T., C. Foussette, and P. Krause, Contemporary evolution strategies. 2013, New York, NY: Springer Berlin Heidelberg. pages cm.

[33] Bäck T., Evolutionary algorithms in theory and practice : evolution strategies, evolutionary programming, genetic algorithms. 1996; New York: Oxford University Press. xii, 314 p.

[34] Hansen N. and A. Ostermeier. Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. in Proceedings of IEEE International Conference on Evolutionary Computation. 1996.

[35] Hansen N. and A. Ostermeier, Completely Derandomized Self-Adaptation in Evolution Strategies. Evolutionary Computation, 2001; 9(2): p. 159-195.

[36] Hansen N. and S. Kern. Evaluating the CMA Evolution Strategy on Multimodal Test Functions. in Parallel Problem Solving from Nature - PPSN VIII. 2004; Berlin, Heidelberg: Springer Berlin Heidelberg.

[37] Kresse G. and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. Computational Materials Science, 1996. 6(1): p. 15-50.

[38] Gale J.D. and A.L. Rohl, The General Utility Lattice Program (GULP). Molecular Simulation, 2003; 29(5): p. 291-341.

[39] Deb K. and H. Jain, An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints. IEEE Transactions on Evolutionary Computation, 2014; 18(4): p. 577-601.